

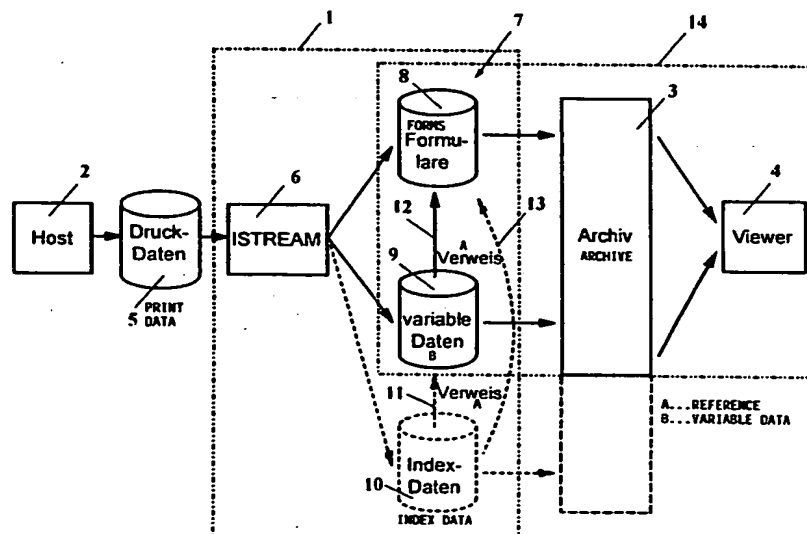


INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation <sup>6</sup> : H04N 1/00	A2	(11) Internationale Veröffentlichungsnummer: WO 99/12337 (43) Internationales Veröffentlichungsdatum: 11. März 1999 (11.03.99)
(21) Internationales Aktenzeichen: PCT/EP98/05460 (22) Internationales Anmeldedatum: 28. August 1998 (28.08.98) (30) Prioritätsdaten: 197 37 811.0 29. August 1997 (29.08.97) DE (71) Anmelder (für alle Bestimmungsstaaten ausser US): OCÉ PRINTING SYSTEMS GMBH [DE/DE]; Siemensallee 2, D-85586 Poing (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US): WEGENER, Harald [DE/DE]; St. Ulrich Strasse 1, D-84149 Velden (DE). (74) Anwälte: SCHAUMBURG, Karl-Heinz usw.; Postfach 86 07 48, D-81634 München (DE).		(81) Bestimmungsstaaten: CA, DE, JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  Veröffentlicht <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>

(54) Title: **METHOD AND DEVICE FOR ELECTRONIC ARCHIVING OF A COMPUTER DATA FLOW**

(54) Bezeichnung: **VERFAHREN UND GERÄT ZUR ELEKTRONISCHEN ARCHIVIERUNG EINES COMPUTER-DATENSTROMS**



(57) Abstract

The invention relates to a method and device for electronic archiving of a data flow (2) which contains graphic and/or text information and is outputted by a computer (2). The print data flow (5) is converted from a printer-specific data format into a pixel based data format (Bit map). Form data (8) is distinguished from variable data (9) in said pixel based data and both categories of data (8,9) are processed differently.

# **(57) Zusammenfassung**

Die Erfindung betrifft ein Verfahren und ein Gerät zur elektronischen Archivierung des von einem Computer (2) abgegebenen Datenstroms (5), der grafische und/oder Textinformationen enthält. Der Druckdatenstrom (5) wird von einem druckerspezifischen Datenformat in ein auf Pixel basierendes Datenformat (Bit Map) umgesetzt. Bei den auf Pixel basierenden Daten werden Formular-Daten (8) von variablen Daten (9) unterschieden und diese beiden Datenarten (8, 9) jeweils unterschiedlich verarbeitet.

## **LEDIGLICH ZUR INFORMATION**

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Letland	SZ	Swasiland
AZ	Aserbaidshjan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

Verfahren und Gerät zur elektronischen Archivierung eines  
Computer-Datenstroms

- 5 Die Erfindung betrifft ein Verfahren und ein Gerät zur elektronischen Archivierung des von einem Computer abgegebenen Datenstroms nach dem Oberbegriff des Patentanspruchs 1.
- 10 Ein solches Verfahren und ein solches Gerät sind aus der DE-4408327 A1 bekannt. In Computern erzeugte Dokumente, die sowohl grafische als auch numerische und alphanumerische Informationen enthalten, werden dabei über eine Schnittstelle, beispielsweise eine Standard-
- 15 Druckerschnittstelle, an ein Archivierungssystem übertragen. Innerhalb des Archivierungssystems werden die zu den Dokumenten gehörenden Daten auf Massenspeichern wie Magnetbändern, magnetischen oder optischen Speicherplatten oder dergleichen langfristig aufbewahrt. Diese Art der
- 20 Speicherung ersetzt in zunehmendem Maße ältere Verfahrensweisen, bei denen auf Papier vorliegende Dokumente bzw. Originale mit einem optischen Abtaster (Scanner) erfaßt, das so erhaltene Bild in elektrische Signale umgewandelt wurden und dann in Archiven abgelegt wurden. Die früher
- 25 übliche Microverfilmung von Dokumenten wird durch diese neue Technik in zunehmendem Maße ersetzt, wenn die Originale bereits in Form von elektronischen Signalen bzw. Computer-Datenströmen vorliegen.
- 30 Elektronische Archivierungssysteme der eingangs genannten Art setzen einen vom Computer ausgegebenen Datenstrom in der Regel in einen Datenstrom mit spezifischem, an das Archiv angepaßten Datenformat um. In vielen Anwendungsfällen ist der Datenstrom des Computers an bestimmte Ausgabesysteme,
- 35 insbesondere an Drucker, angepaßt. Beispiele solcher

Druckdatenströme sind das von IBM geprägte Format IPDS oder das von Hewlett Packard geprägte Datenformat PCL.

Ein Archivierungssystem setzt diese Datenströme auf ein  
5 Format um, welches dem Archivierungssystem entspricht. Die direkte Speicherung der vom Computersystem abgegebenen Daten erweist sich dabei als sehr nachteilig, weil dabei eine sehr große Anzahl an Systemparametern, z.B. Zeichenfonts, mit abgespeichert werden müßte. Auch die Reproduktion derart  
10 abgespeicherter Daten würde sich dann als sehr aufwendig erweisen. Archivierungssysteme speichern die Daten deshalb auf einer Pixel-orientierten Basis, z.B. im sog. TIF-Format, ab. Durch eine derartige, punktweise aufgebaute Speicherungstechnik eröffnet sich dann die Möglichkeit, das  
15 Datenvolumen nach gängigen Komprimierungsverfahren zu verringern. Das Datenvolumen derart komprimierter Daten ist um so größer, je mehr schwarz-weiß Übergänge zu verarbeiten sind. Andererseits werden Druckseiten in zunehmendem Maße optisch immer aufwendiger gestaltet. Beispielsweise werden  
20 auf Formularen immer häufiger Grauraster hinterlegt, um diese optisch ansprechender zu gestalten und Hervorhebungen besser kenntlich zu machen. Diese Entwicklung hat zur Folge, daß bei der Archivierung Druckseiten einen immer größeren Speicherbedarf haben.

25 In einem anderen, bekannten System, dem sog. COLD-System, werden Daten zur Archivierung getrennt als überwiegend grafische Daten und als überwiegend kodierte Daten (Zeilendaten) abgelegt. Bei erweiterten COLD-Verfahren werden  
30 Rohdaten und Ressourcen ebenfalls getrennt abgelegt und bei der Reproduktion wird der gesamte Druckprozess nachgebildet. Dies bedingt eine komplexe Ressourcen-Verwaltung.

Aus der DE 195 15 981 A1 ist ein Verfahren zum Erfassen  
35 manuell beschriebener Belege bekannt, bei dem die Belege

gescannt und anschließend unter Ausblendung von Vordruckinformationen an Bildschirmen weiterbearbeitet werden. Da die Vordruckinformationen bei der Weiterverarbeitung nicht mehr verfügbar sind, eignet sich dieses Verfahren nur dann, wenn die auf der Vorlage abgedruckten Informationen zum Zeitpunkt der Weiterverarbeitung noch bekannt bzw. verfügbar ist. Für ein langfristiges Archivierungssystem ist dieses Verfahren deshalb kaum geeignet.

In der Veröffentlichung Wong, K.Y. et. al., „Document Analysis System“, in IBM J. Res. Develop., Vol. 26, No. 6, Nov. 1982, Seiten 647-656 ist ein Verfahren zur Unterscheidung zwischen Textdaten und graphischen Daten beschrieben. Es eignet sich zum manuellen Bearbeiten von gescannten Dokumenten, ist aber für die Archivierung von Druckdaten nicht ohne weiteres einsetzbar.

Aufgabe der Erfindung ist es, ein System zur Archivierung von Computer-Datenströmen bereit zu stellen, bei dem ein hoher Datenkompressionsgrad erzielt werden kann, auch wenn aufwendige grafische Informationen archiviert werden sollen.

Diese Aufgabe wird durch die im Patentanspruch 1 und im Patentanspruch 15 beschriebene Erfindung gelöst. Vorteilhafte Ausführungsformen der Erfindung sind Gegenstand der Unteransprüche.

Die Erfindung beruht auf der Erkenntnis, daß ein Großteil der zu archivierenden Computer-Datenströme, insbesondere an Drucker zugeführte Datenströme, aus zweierlei Datenarten besteht. Einerseits sind in diesen Druckdaten häufig Formulare hinterlegt, die für eine bestimmte Anzahl zusammengehöriger Druckdatensätze, sog. Jobs, in ein und derselben Weise wiederkehren. Andererseits sind diese

Formulare mit variablen Daten gefüllt, welche jeweils das Charakteristische eines Dokuments darstellen. Erfindungsgemäß wurde erkannt, daß der durch Formulare hervorgerufene Anteil an komprimiertem Druck-Datenstrom oft denjenigen Anteil  
5 überwiegt, der durch die variablen Daten hervorgerufen wird. Erfindungsgemäß werden deshalb diejenigen Daten, die von Formularen stammen, von denjenigen Daten, die variabel sind, getrennt, und die beiden Datenarten jeweils getrennt auf Bit-Map-Basis, insbesondere komprimiert, weiterverarbeitet.  
10 Innerhalb des Archivierungssystems werden diese Daten voneinander getrennt abgespeichert und später zum Betrachten wieder zusammengeführt; dies kann insbesondere Pixel-genau erfolgen. Für eine Gruppe zusammengehörender Daten (Job) werden die Daten eines Formulars nur einmal abgelegt. Bei den  
15 einzelnen Dokumenten werden die zugehörigen variablen Daten mit dem Formular nur noch in der Weise verknüpft, daß ein Verweis auf den entsprechenden Speicherbereich des Formulars abgelegt wird. Vorteilhaft ist dabei, daß sich die - in der Regel umfangreiche Datenmenge eines Formulars - für jeden Job  
20 auf ein einmaliges Abspeichern des Formulars beschränkt. Hierdurch wird eine beträchtliche Speicherplatzeinsparung gegenüber Systemen des Standes der Technik erzielt.

Bei Druckgut mit hoher Auflage, d.h. mit häufig  
25 wiederkehrenden, gleichartigen Formularen, wird hierdurch ein hoher Grad an Speicherplatzeinsparung erreicht.

Vorteilhaft ist weiter, - insbesondere gegenüber dem o.g. COLD-Systemen - daß sich die Reproduktion der Daten sehr  
30 einfach gestaltet, weil auf Bit-Map-Basis (z.B. TIF-Format) nur ein Viewer zur Reproduktion benötigt wird (selbsttragendes Format).

Als vorteilhaft erweist es sich auch, zusammen mit den  
35 Verweisen eine Information abzuspeichern, die ein genaues

Zusammenführen der Formulardaten mit den zugehörigen, variablen Daten ermöglicht.

- In einer weiteren, vorteilhaften Ausführungsform der Erfindung werden die Formular-Daten je Job, d.h. je zusammenhängendem Datenstrom, nur einmal im Archivspeicher als Bit Map abgespeichert, während die variablen Daten jedes Dokuments des Jobs jeweils einzeln abgespeichert werden.
- 10 In einer auf extrem hohen Komprimierungsgrad ausgerichteten Ausführungsform werden die Formulardaten inhaltlich reduziert oder sogar gänzlich unterdrückt. In einem derartigen System würden dann im Archivspeicher nur noch die variablen Daten abgespeichert.
- 15 Eine Unterscheidung zwischen Formulardaten und variablen Daten erfolgt vorzugsweise im druckerspezifischen Datenformat, d.h. bevor der Druckdatenstrom bzw. die betroffenen Teile davon in das pixelbasierte Datenformat umgesetzt werden. Dabei ist es insbesondere günstig, Formularindikatoren zur Erkennung von Formular-Daten zu verwenden wie beispielsweise die in IPDS-Druckdatenströmen üblichen Overlay-Daten oder in PCL-Druckdatenströmen Makros, formularspezifische Grafiken wie grau hinterlegte Rechtecke oder dergleichen. Dabei muß nicht zwangsläufig ein gesamtes, einer Seite entsprechendes Formular gesucht und gefunden werden. Erfindungsgemäß kann bereits durch das Erkennen einzelner Formularbestandteile ein hoher Grad an Speicherplatzeinsparung erzielt werden. Auf jeder Seite können mehrere Bestandteile auftreten.
- 20
- 25
- 30

- Werden Formularindikatoren verwendet, so kann eine hohe Performance beim Archivieren erreicht werden, weil Formulardaten bzw. ,-strukturen relativ einfach bzw. eindeutig erkennbar sind. Dabei kann es vorteilhaft sein,
- 35

Formularindikatoren relativ restriktiv einzusetzen, wobei  
Formulardaten nur dann als solche bezeichnet werden, wenn mit  
relativ einfachen Beurteilungskriterien sicher ist, daß die  
entsprechenden auf Pixelbasis umgesetzten Daten pixelgenau  
5 übereinstimmen. Zur Performancesteigerung während des  
Archivierungsprozesses kann auch beitragen, wenn für die zu  
erkennenden Formulardaten eine Mindestgröße vorgesehen ist.  
Sehr kleine Bildelemente in einem Druckdatenstrom werden dann  
nicht der Prozedur zur separaten Rasterung und einer  
10 Zwischenspeicherung zugeführt sondern gemeinsam mit den  
variablen Daten gerastert und gespeichert.

Bei der Suche der Daten in dem Archivspeicher bzw. bei der  
Rekonstruktion des ursprünglichen Datenstroms können die  
15 Formulardaten mit den Inhaltsdaten entweder automatisch  
überlagert werden anhand von Steuerungssignalen (Fig. 1,  
Verweis) oder aber variable Daten und Formulardaten getrennt  
voneinander geladen und vom Bediener gesteuert übereinander  
gelegt werden.

20 Im Zuge des Archivierens können neben den variablen Daten und  
den Formulardaten sog. Index-Daten erzeugt werden, in denen  
Suchkriterien abgespeichert werden, sowie die Verweisdaten  
auf die Speicherorte der variablen Daten und gegebenenfalls  
25 der Formulardaten.

In einer weiteren, vorteilhaften Ausführungsform der  
Erfindung erfolgt der Datenübertrag zum Computersystem in das  
Archivsystem über einen Drucker-Controller, welcher mit einem  
30 nachgeschalteten Computer zusammenarbeitet. Der Druck-  
Controller kann dabei also sowohl zum Erzeugen eines  
Pixeldatenstroms zur Ansteuerung eines Druckers, als auch zum  
Erzeugen eines Archiv-Datenstroms verwendet werden. Die  
beiden Datenstromverbindungen können insbesondere zeitlich  
35 parallel erfolgen, wodurch eine hohe Geschwindigkeit des



Gesamtprozesses (Drucken und Archivieren) erreichbar ist. Vom Controller zum Computer erfolgt der Datentransfer sowohl für die Formular- und variablen Daten des punktorientierten Datenformats (Bit Map) als auch die im zeichenorientierten (ASCII-Format) Indexdaten über eine gemeinsame Schnittstelle.

Die Verknüpfung, d.h. der Verweis von den variablen Daten zu den Formulardaten, kann innerhalb des variablen Datensatzes hinterlegt werden, innerhalb des entsprechenden Index-Datensatzes oder in beiden Datensätzen.

Weitere Vorteile und Wirkungen der Erfindung werden anhand der nachfolgend beschriebenen Ausführungsbeispiele deutlich, die im Zusammenhang mit zwei Figuren beschrieben werden.

Es zeigen:

**Figur 1:** ein elektronisches Archivierungssystem,

**Figur 2:** elektronische Komponenten des Archivierungssystems,

**Figur 3:** einen Ablauf zum Erkennen und Trennen der verschiedenen Datentypen in einem Druckdatenstrom,

**Figur 4:** einen Ablauf zum Archivieren von getrennt abgespeicherten Daten und

**Figur 5:** einen Ablauf zum Drucken eines vorab getrennten Druckdatenstroms.

Bei dem in Figur 1 dargestellten elektronischen Archivierungssystem wird der von einem Host-Computer 2 abgegebene Datenstrom über eine Archivierungs-Schnittstelle 1 an einen Archiv-Speicher 3 abgegeben. Der Archiv-Speicher 3

kann insbesondere ein Gerät zum Beschreiben von Laser-Speicherplatten sein. Die im Archivspeicher 3 gespeicherten Daten können in dem Bit-Map-basierten Datenformat mit einem Lesegerät (Viewer) 4 abgerufen und beispielsweise wiederum  
5 auf einem Computerbildschirm dargestellt werden.

Der vom Host-Computer abgegebene Datenstrom 2 wird über eine Standard-Druckerschnittstelle in einen Controller 6 (ISTREAM) eingespielt. Diese Daten sind in Form sog. Jobs  
10 gruppiert. Ein Job ist ein zusammenhängender Datenstrom, der in irgendeiner Weise (logisch oder physikalisch) miteinander verknüpft ist. Innerhalb des Host-Computers kann beispielsweise eine Marke von außen zugeführt werden, die bestimmte Druck-Datenströme miteinander zu einem Job  
15 verbindet.

Innerhalb der Archivierungs-Schnittstelle werden die vom Host-Computer gelieferten Daten 5 klassifiziert nach variablen Daten 9 und Formular-Daten 8. Die Formular-Daten  
20 (z.B. Overlays) stellen dabei einen fixen Anteil des Jobs dar, der auf mehreren Druckseiten des Jobs unverändert auftritt. Die variablen Daten sind dagegen innerhalb eines Jobs von Druckseite zu Druckseite unterschiedlich.

25 In der Archivierungs-Schnittstelle werden nunmehr für jeden Job die Formulardaten 8 für jedes Formular nur jeweils einmal hinterlegt, während die variablen Daten seitenweise hinterlegt werden. Gleichzeitig werden Index-Daten erzeugt, die zu den variablen Daten bestimmte Zusatzinformation  
30 enthalten, beispielsweise herausgegriffene Ordnungsnummern, Namen, Stichworte oder dergleichen. Die Index-Daten werden im Zuge der weiteren Verarbeitung in Form einer Datenbank weiterverarbeitet und ermöglichen dem späteren Benutzer des Archivs, bestimmte Datensätze wiederzufinden. Die Index-Daten  
35 enthalten für alle variablen Datensätze Verweis-Informationen

11, welche die Zuordnung bzw. das Wiederfinden der variablen Daten 9 ermöglichen.

Um alle variablen Daten einer Druckseite auch mit dem zugrundeliegenden Formular-Datensatz 8 zu verbinden, muß ein Verweis zwischen variablen Daten und Formular erfolgen. Dieser kann entweder direkt als Verweis 12 oder indirekt über den Index-Datensatz 10 als indirekter Verweis 13 erfolgen. Zur Sicherheit können auch alle drei Verweisarten 11, 12 und 13 gemeinsam für jede Druckseite hinterlegt werden. Die so erzeugte Datenstruktur 7 wird dann über eine geeignete Verbindung dem Archiv-Speicher 3 zugeführt.

Die Formular-Daten 8 und die variablen Daten 9 sind - genauso wie die Archiv-Daten - innerhalb einer Bit Map-Datendomäne 14, während die Index-Daten in einem kodierten Format, z.B. ASCII vorliegen.

Figur 2 zeigt eine Ausführungsform für eine erfindungsgemäße Archivierungs-Schnittstelle 1. Sie ist unterteilt in einen Drucker-Controller 6 und einen nachgeschalteten Computer 15, welcher beispielsweise ein Personal Computer (PC) sein kann.

Die vom Host-Computer 2 gelieferten Daten werden über einen Kanalanschluß 16 eingespeist. Von diesem werden sie in einem Zwischenspeicher 17 zwischengespeichert. Aus diesem Speicher werden die Daten sukzessive von einer Rastersteuerung 18 entnommen, die die Druckdaten in ein Bit Map Muster umsetzt. Die Rastersteuerung 18 verfügt hierzu über einen nicht näher gezeigten Speicher, welcher Rasterinformationen zur Umsetzung der Druckdatensprache in die entsprechenden Bit Map Informationen enthält. Die Rastersteuerung 18 unterscheidet innerhalb eines zusammengehörenden Druckjobs Formulardaten von variablen Daten. Gleichzeitig produziert sie zu jedem Job einen Index-Datensatz, welcher zur Verwaltung der zu

archivierenden Daten dient. Die Rastersteuerung 18 speichert die variablen Daten in einem variablen Datenspeicher 19, die Formulardaten in einem Formulardatenspeicher 20 und die Index-Daten in einem Index-Datenspeicher 21 ab. Diese drei Speicher 19, 20 und 21 sind Bestandteil des flüchtigen Arbeits-Bildspeichers 22, des Controllers 6. Die in den Datenspeicher 19, 20 und 21 erzeugten Daten werden dann über eine gemeinsame Schnittstelle 23 an den nachgeschalteten Personal Computer 15 übertragen. Den Datentransfer steuert eine Gesamtsteuerung 24 des Controllers 6, welche mit einer Bedienoberfläche und Steuerung 25 des Personal Computers 15 verbunden ist.

Die Rastersteuerung 18 behandelt Formular-Daten und variable Daten unterschiedlich. Für Formulardaten werden eigene Bit Maps im Formularspeicher erzeugt. Für die variablen Daten werden im Speicher 19 Bit Maps erzeugt, die keine Formular-Daten (Overlays) enthalten.

Die Index-Daten, welche im Index-Speicher 21 abgelegt werden, enthalten Verwaltungsinformationen, welche u.a. für jede Seite des Druckjobs die verwendeten Formular-Overlays ablegt. Dabei wird auch der Bezug zwischen variablen Daten und Formular-Daten hergestellt.

Innerhalb des Personal Computers 15 werden die aus dem Speicher für variable Daten 19 und dem Speicher für Formulardaten 20 stammenden Daten in einer Komprimierungsstufe 26 komprimiert. Die Index-Daten werden vorzugsweise ohne Kompression in einer Index-Warteschlange 27 zwischengespeichert, bis sowohl variable als auch Formulardaten die Komprimierung durchlaufen haben. Im Anschluß an die Komprimierungsstufe 26 sind zwei parallel geschaltete Warteschlangen 28, 29 für die variablen Daten und für die Formular-Daten vorgesehen, in denen die

## 11

entsprechenden Daten so lange zwischengespeichert werden, bis die zugehörigen übrigen Daten die Komprimierungsstufe ebenfalls durchlaufen haben.

- 5 Die in den Warteschlangen 27, 28 und 29 befindlichen Daten werden dann von einer Synchronisierungsstufe 30 gemeinsam entnommen und über eine Netzwerksteuerung 31 zur Ausgabeschnittstelle 32 in Richtung zum Archivspeicher 3 abgegeben.

10

In einer weiteren vorteilhaften Ausführungsform können alle Verarbeitungseinheiten des Personal-Computers 15 auch in dem Drucker-Controller integriert sein.

- 15 In Figur 3 ist ein Ablaufdiagramm dargestellt, das einen Ablauf zur Analyse von Druckdaten darstellt. In einem ersten Schritt S1 wird jeweils eine Gruppe von Druckdaten aus dem Druckdatenstrom 5 in einen vorgegebenen Bereich des Controllers 6 eingelesen. Der Umfang der Gruppe kann
- 20 individuell bestimmt werden, wobei eine Druckseite beispielsweise in 2, 5, 10 oder noch mehr Teilbereiche unterteilt sein kann. Die nachfolgende Analyse der Druckdaten erfolgt im sog. CI-Raum (Coded Information Raum). Die Analyse erfolgt dabei so, daß die Auswirkungen der Datenumsetzung auf
- 25 Pixel-basierter Basis, d.h. im sog. NCI-Raum (Non-Coded Information Raum) berücksichtigt werden. Dazu wird der Druckdatenstrom bzw. die Gruppen der Druckdaten nach gewissen Kategorien auf Formularindikatoren untersucht. Ein erster Formularindikator, der beispielsweise in Druckdatenströmen
- 30 nach dem IPDS-Format häufig verwendet wird, sind sogenannte Overlay-Daten oder Overlay-Aufrufe. Liegen solche Overlay-Daten oder ein Overlay-Aufruf vor, so ist dies ein Hinweis darauf, daß den Daten ein Formular zugrunde liegt. Die Overlay-Daten bzw. der Overlay-Aufruf werden dann gemäß
- 35 Schritt S7 weiterverarbeitet, der weiter unten noch näher

beschrieben wird. Bei Overlays und Makros wird sinnvollerweise der jeweilige Aufruf ausgewertet, da ein definiertes Overlay oder Makro nicht notwendigerweise auch verwendet wird. Wurde ein Overlay also bereits einmal  
5 aufgerufen, so wird bei jedem späteren Aufruf dieses Overlays im Schritt S7 der Status „ja“ ausgegeben.

Die Untersuchung der Druckdaten kann relativ restriktiv erfolgen, d.h. Formularindikatoren nur dann als gegeben  
10 betrachtet werden, wenn sie relativ eindeutig darauf hinweisen, daß ein Formular vorliegt, d.h. daß die entsprechenden Daten innerhalb eines Druckjobs, der eine Vielzahl von Seiten enthält, mehrmals vorkommen.

15 Wurde im Schritt S2 kein Overlay festgestellt, so wird im Schritt S3 untersucht, ob in der aktuellen Gruppe von Druckdaten Makro-Informationen vorhanden sind, ggf. wird das jeweilige Makro hinsichtlich typischen Makro-Formularindikatoren im Schritt S13 überprüft. In der  
20 Drucksprache PCL werden häufig wiederkehrende Aktionen regelmäßig in Makros abgelegt. Somit ist die Wahrscheinlichkeit, bei der Untersuchung von Makros auf konstante Daten mit Formelcharakter, beispielsweise auf Felder mit grau hinterlegtem Hintergrund zu stoßen.

25 Gegebenenfalls wird im Schritt S6 eine Variable, die den Formelindikator repräsentiert, auf 1 (ja) gesetzt und die entsprechende Druckdatengruppe im Schritt S7 weiterbehandelt.

Wie im Schritt S3 werden Daten, die kein Makro enthalten,  
30 auch im Schritt S4 hinsichtlich graphischen Objekten untersucht. Auch hier sind beispielsweise mit Graurastern gefüllte Flächen im Schritt S14 erkennbar und wiederum mit dem Schritt S6 als Formularbestandteil zu kennzeichnen. Insbesondere werden gefüllte Rechtecke meist durch einfach zu  
35 erkennende Strukturen in entsprechenden Drucksprachen

beschrieben. Auch kleinere Bit-Maps in bestimmten Bereichen, z.B. ein Firmenlogo, auf den Druckseiten können lohnenswerte Objekte sein, die als Formularbestandteil gekennzeichnet werden können.

5

Ergibt auch Schritt S4 kein Formular-Ergebnis, so kann in einem Schritt S5 geprüft werden, ob in den aktuellen Druckdaten typische, häufig wiederkehrende Textbestandteile vorhanden sind. Bei dieser Prüfung kann u.U. restriktiver vorgegangen werden, wenn die Textdaten zu einem hohen Grad individualisiert sind. Sind jedoch solche Textdaten eindeutig als Formularbestandteile in Schritt S15 ermittelt, so wird auch hier der Schritt S6 zum Schritt S7 führen.

10

15 Textobjekte sollten ebenfalls nur in bestimmten Bereichen einer Seite, z.B. in der Fußzeile auf konstante Anteile untersucht werden. Dabei können beispielsweise auch Schriftattribute untersucht werden. Beispielsweise deutet eine Schriftgröße < 10 pt darauf hin, daß eine Information  
20 eine Formularinformation ist.

Wurde im Schritt S6 ein Formularindikator entdeckt, so wird im Schritt S7 geprüft, ob diese (im NCI-Raum identische) Struktur bereits früher aufgetreten ist. Wenn diese Bedingung  
25 nicht erfüllt ist, so wird im Schritt S17 die aktuell aufgetretene Struktur in einem Strukturspeicher 40 abgespeichert. Die Daten werden sodann im Rahmen des normalen Rasterprozesses im Schritt S9 in der Bit-Map 19a für variable Daten im Rahmen des normalen Rasterprozesses S9 abgelegt.

30

Wird im Schritt S7 dagegen festgestellt, daß im Speicher 40 bereits eine Struktur gespeichert ist, die den aktuellen Daten entspricht, so wird im Schritt S18 geprüft, ob die konstanten Daten bereits bei einer vorhergehenden Seite  
35 aufgerastert wurden und somit schon als Bitmap vorliegen. In

diesem Fall können im Schritt S12 direkt Referenzen zwischen den in den Bitmaps 20a, 20b, 20c gespeicherten konstanten Daten und den jeweils zugehörigen variablen Daten gebildet werden. Bei dieser Referenzbildung wird der Speicherort der  
5 Bitmaps im Festspeicher 20 und der Ort der Bitmap bezüglich der Druckseite (sog. xy-Position) in einer Referenztabelle 41 festgehalten. Falls die konstanten Daten noch nicht aufgerastert sind, werden diese Daten in einem eigenen Rasterprozeß für konstante Daten im Schritt S8 gerastert und  
10 in einer Bit-Map für konstante Daten jeweils separat als Bit-Map 20a, 20b, 20c usw. abgespeichert.

Nachdem die Schritte S18, S8, S9 und S12 abgeschlossen sind, wird im Schritt S10 überprüft, ob die aktuellen Daten ein  
15 Seitenende markieren; wenn nicht, werden mit Schritt S1 die nächsten Druckdaten eingelesen. Ist das Seitenende erreicht, so wird im Schritt S11 eine logische Statusvariable auf einen Wert „TRUE“ gesetzt.

20 Beim erstmaligen Auftreten einer Struktur mit Formularindikator wird diese Struktur lediglich im Coded-Format im Strukturspeicher 40 abgespeichert. Durch dieses Verfahren werden nur Strukturen, die wiederholt vorkommen, einer Sonderbehandlung unterworfen. Beim erstmaligen  
25 Auftreten einer solchen Struktur werden die entsprechenden Daten zunächst immer als variable Daten interpretiert.

Bei bestimmten Ereignissen bzw. Kommandos innerhalb des Druckdatenstroms (z.B. beim Beginn eines neuen Druckjobs, bei  
30 einem Printer-Reset oder dergl.) werden die aufgetretenen Strukturen im Strukturspeicher 40 gelöscht.

Beim wiederholten Auftreten einer im NCI-Raum identischen Struktur wird dieses im Schritt S12 in der Referenztabelle 41  
35 vermerkt.



## 15

- In Figur 4 ist der Ablauf zum Archivieren der zuvor gemäß Figur 3 verarbeiteten Druckdaten beschrieben. Im Schritt S20 wird zunächst geprüft, ob die Seitenende-Statusvariable (vgl. Schritt S11), auf „TRUE“ gesetzt ist. Ist diese Bedingung
- 5 erfüllt, so werden die im Schritt S9 bzw. in Bit-Map 19a gespeicherten, variablen Daten im Schritt S21 komprimiert und danach im Schritt S22 Verweise zu konstanten Daten anhand der Referenztabelle 41 gebildet und abgelegt.
- 10 Im Schritt S23 wird dann geprüft, ob zu den in Schritt S22 gebildeten Verweisen bereits konstante Daten im Archivspeicher abgelegt sind. Wenn dies nicht der Fall ist, so werden die entsprechenden konstanten Daten aus den Bit-
- 15 Maps 20a, 20b, 20c usw. unter Verwendung der Referenztabelle 41 im Schritt S24 komprimiert und die komprimierten Daten im Schritt S25 im Archivspeicher abgelegt.

- Wird dagegen im Schritt S23 festgestellt, daß im Archivspeicher bereits entsprechende Daten komprimiert
- 20 abgelegt sind, so werden im Schritt S26 nur noch die aktuellen variablen Daten komprimiert und im Archivspeicher abgelegt. Danach wird im Schritt S20 der nächste Datensatz behandelt, bis alle ankommenden Daten verarbeitet sind.
- 25 Die variablen Daten werden in einem separaten Speicher aufgerastert und komprimiert. Beim Abspeichern der variablen Daten werden die Verweise, insbesondere diejenigen auf den XY-Offset innerhalb der Seite sowie diejenigen auf den Speicherort, der auf der Seite enthaltenen konstanten Daten
- 30 hinzugefügt.

- In Figur 5 ist der Prozeß zum Druckern der gemäß Figur 3 verarbeiteten Daten dargestellt. Im Schritt S30 wird wiederum überprüft, ob alle Daten einer Seite abgearbeitet sind, d.h.
- 35 ob in Schritt S11 von Figur 3 die Variable zum Seitenabschluß

auf „TRUE“ gesetzt ist. Ist dies der Fall, so kann zum Drucken in Schritt S31 die in Speicher 19 gespeicherte Bit-Map 19a der variablen Daten einer Seite mit den entsprechenden Bit-Maps 20a, 20b, 20c der konstanten Daten der zugehörigen Seite anhand der Referenztabelle 41 überlagert werden. Dies kann beispielsweise durch eine logische „ODER“-Operation zwischen dem entsprechenden Bit-Map geschehen. Anschließend kann in Schritt S32 der Druckvorgang gestartet werden.

10

Ausgehend von dem in Figur 3 dargestellten Verfahren zur Untersuchung der Daten können die Prozesse der Figuren 4 und 5, d.h. Archivieren und Drucken, parallel erfolgen.

15

Es wurden Ausführungsbeispiele der Erfindung beschrieben. Dabei ist klar, daß ohne weiteres Abwandlungen und Weiterentwicklungen der Erfindung angegeben werden können. Beispielsweise kann in einem Druck- und Archiviersystem vorgesehen sein, daß der Anwender bzw. das Anwendungsprogramm bereits beim Erzeugen des Datenstroms immer wiederkehrende Daten (Formulardaten) mit besonderen Kennzeichnungen versieht, (z.B. durch eine entsprechende Steuerinformation) und daß diese Kennzeichnungen im späteren Archivier- bzw. Druckvorgang entsprechend ausgewertet werden. Ähnlich wie bei der Overlay-Information kann damit eine sichere Entscheidung getroffen werden, daß bestimmte Daten Formulardaten sind.

20

25

30

Es kann auch vorteilhaft sein, im Datenaufbereitungsprozeß (Fig. 3) nach dem Rastern der konstanten Daten im Schritt S8 gleich die Schritte S24 und S25 durchzuführen und die konstanten Daten im Archivspeicher abzulegen. Hierzu müssen aber die Ablagearten der Daten frühzeitig bekannt sein.

## Patentansprüche

1. Verfahren zur elektronischen Archivierung des von einem  
5 Computer (2) abgegebenen Datenstroms (5), der grafische  
und/oder Textinformationen enthält, wobei der Datenstrom(5)  
von einem druckerspezifischen Datenformat in ein auf Pixel  
basierendes Datenformat (Bit Map) umgesetzt wird, dadurch  
gekennzeichnet, daß bei den auf Pixel basierenden Daten  
10 Formular-Daten (8) von variablen Daten (9) unterschieden  
werden und daß diese beiden Datenarten (8, 9) jeweils  
unterschiedlich verarbeitet werden.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß den  
15 variablen Daten (9) Verweise zu den Formulardaten (8)  
zugeordnet werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet,  
daß innerhalb einer vorgegebenen Datengruppe (Job) ein  
20 Formular-Datensatz gleichartiger Formulardaten nur einmal  
abgespeichert wird, während die zugeordneten variablen  
Daten aller Datensätze der Datengruppe (Job) jeweils alle  
gespeichert werden.
- 25 4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch  
gekennzeichnet, daß eine Unterscheidung zwischen  
Formulardaten (8) und variablen Daten (9) im  
druckerspezifischen Datenformat erfolgt.
- 30 5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß im  
Datenstrom (5) Formularindikatoren zur Erkennung von  
Formular-Daten (8) gesucht werden.
- 35 6. Verfahren nach Anspruch 4 oder 5, dadurch gekennzeichnet,  
daß die Daten des Datenstroms (5) gruppenweise zunächst auf

Formularindikatoren untersucht werden und nur bei wiederholtem Auftreten von Formulardaten (8) die Zuordnung zwischen den variablen Daten (9) und den Formulardaten (8) erfolgt.

5

7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß als Formularindikatoren Overlay-Informationen, besondere Steuerinformationen, Makro-Informationen, graphische Informationen, vorgegebene Textbausteine und/oder vorgegebene Textattribute verwendet werden.

10

8. Verfahren nach einem der Ansprüche 4 bis 7, dadurch gekennzeichnet, daß ein Formular-Datensatz nach dem ersten Auftreten innerhalb der vorgegebenen Datengruppe (Job) des Druckdatenstroms (5) abgespeichert wird und erst nach wiederholtem, insbesondere zweitem Auftreten als Formulardatensatz markiert, in eine Formular-Bit-Map (20a, 20b, 20c) umgesetzt und dem zugehörigen variablen Datensatz (9) zugeordnet wird.

15

20

9. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß innerhalb eines Arbeitsablaufes wahlweise entweder gedruckt oder archiviert, oder gleichzeitig gedruckt und archiviert wird.

25

10. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Formulardaten im Archivspeicher (3) nicht abgespeichert werden.

30

11. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß das ursprüngliche Pixelbild aus den Formular-Daten (8) und den variablen Daten (9) rekonstruiert wird.

12. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß unter Verwendung der Verweise eine Überlagerung der Formular-Daten (8) und der variablen Daten (9) erfolgt.

5

13. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß ein Index-Datensatz erzeugt wird.

10

14. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß der Index-Datensatz einen Verweis auf die variablen Daten (9), insbesondere auf die Formular-Daten (8), enthält.

15

15. Gerät zur elektronischen Archivierung des von einem Computer abgegebenen Datenstroms (5), der grafische und/oder Textinformationen enthält, wobei der Druckdatenstrom (5) von einem druckerspezifischen Datenformat in ein auf Pixel basierendes Datenformat (Bit Map) umgesetzt wird, dadurch gekennzeichnet, daß eine Archivierungs-Schnittstelle (1) vorgesehen ist, die im auf Pixel basierenden Datenformat Formular-Daten (8) und variable Daten (9) unterschiedlich verarbeitet.

20

16. Gerät nach Anspruch 15, gekennzeichnet durch einen Drucker-Controller (6), der über eine Schnittstelle variable Daten (9), Formular-Daten (8) und Index-Daten (10) an einen Weiterverarbeitungscomputer (PC) übergibt.

25

17. Gerät nach Anspruch 16, bei dem die Verarbeitungseinheiten des Weiterverarbeitungscomputers (PC) in den Drucker-Controller integriert sind.

30

18. Gerät nach Anspruch 15 bis 17, dadurch gekennzeichnet, daß in der Archivierungs-Schnittstelle (1) Formular-Daten

(8) und variablen Daten (9) unterschieden werden.

19. Gerät nach einem der Ansprüche 15 bis 18 dadurch gekennzeichnet, daß zur Unterscheidung zwischen  
5      Formulardaten (8) und variablen Daten (9) der Datenstrom  
    (5) im druckerspezifischen Datenformat untersucht wird.

1 / 7

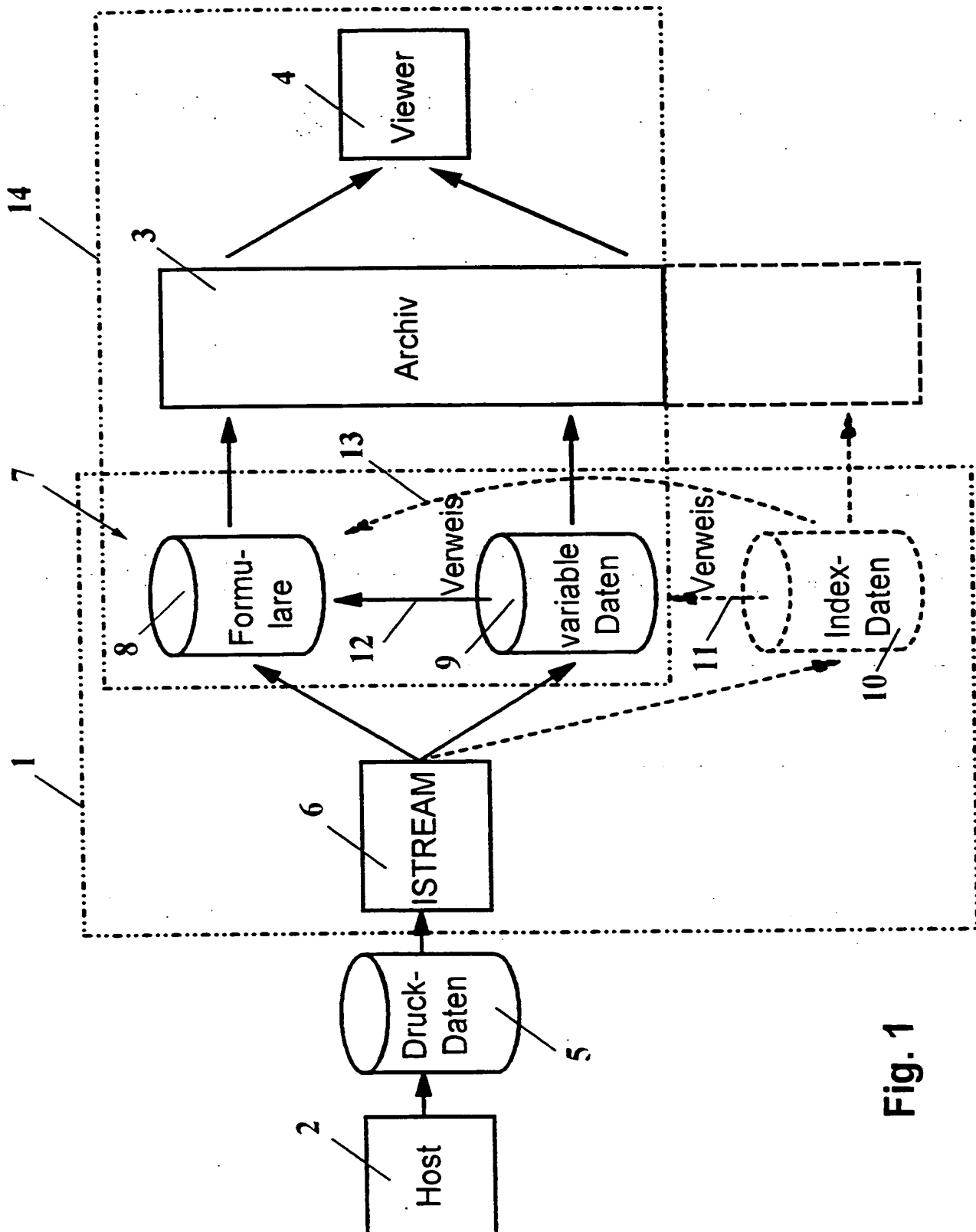
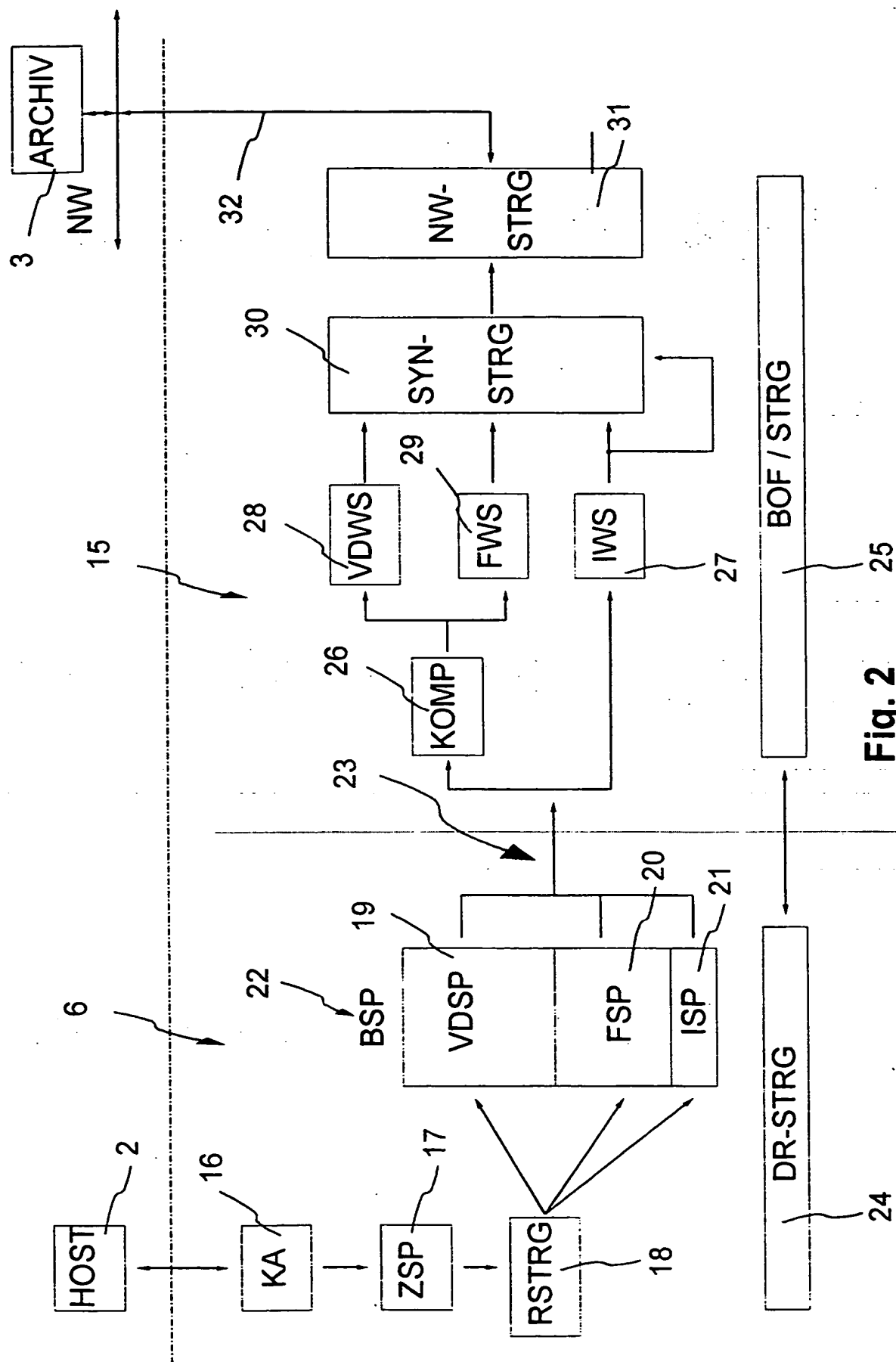


Fig. 1

**THIS PAGE BLANK (USPTO)**

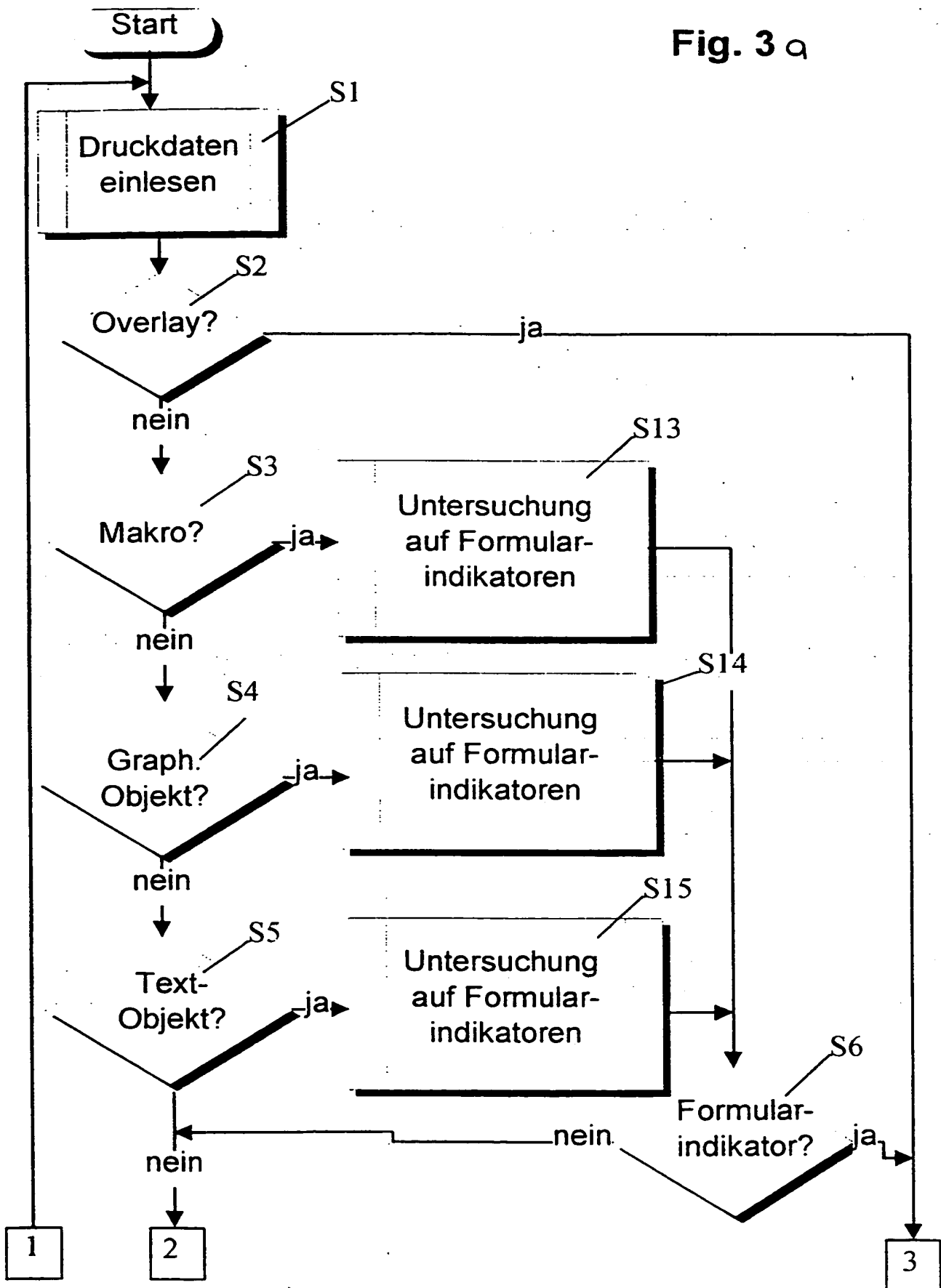




**Fig. 2**

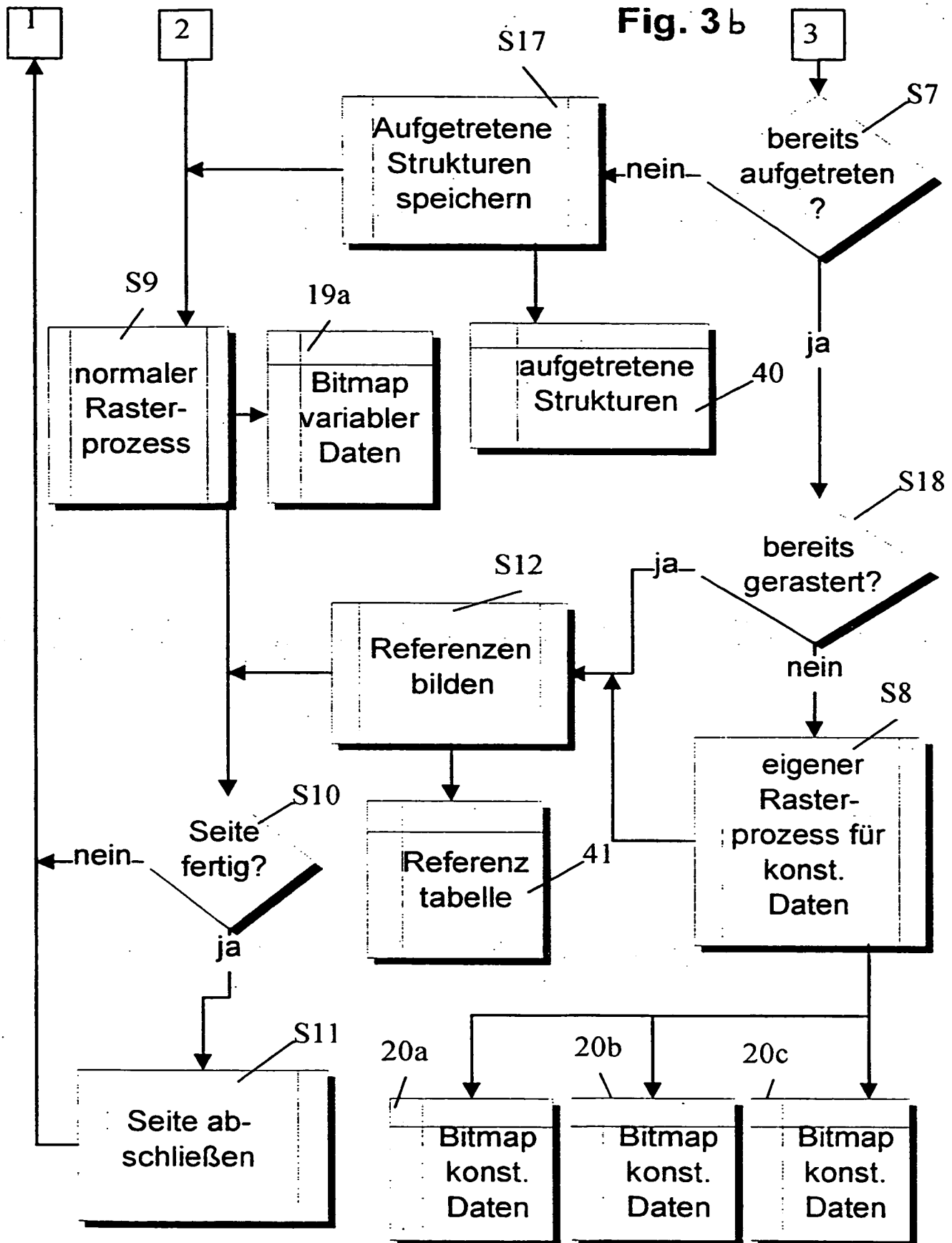
**THIS PAGE BLANK (USPTO)**

Fig. 3 a



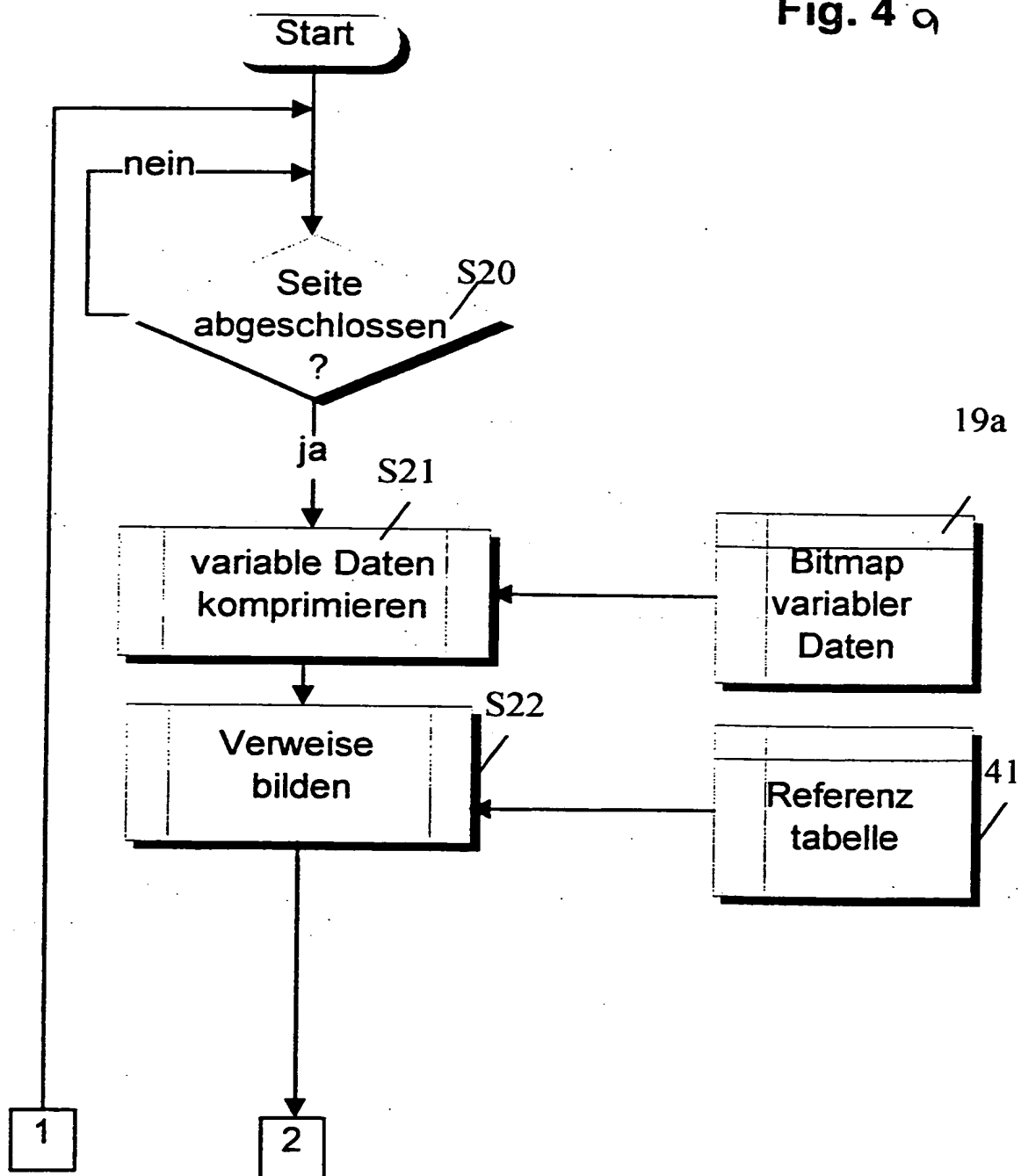
**THIS PAGE BLANK (USPTO)**

Fig. 3 b



**THIS PAGE BLANK (USPTO)**

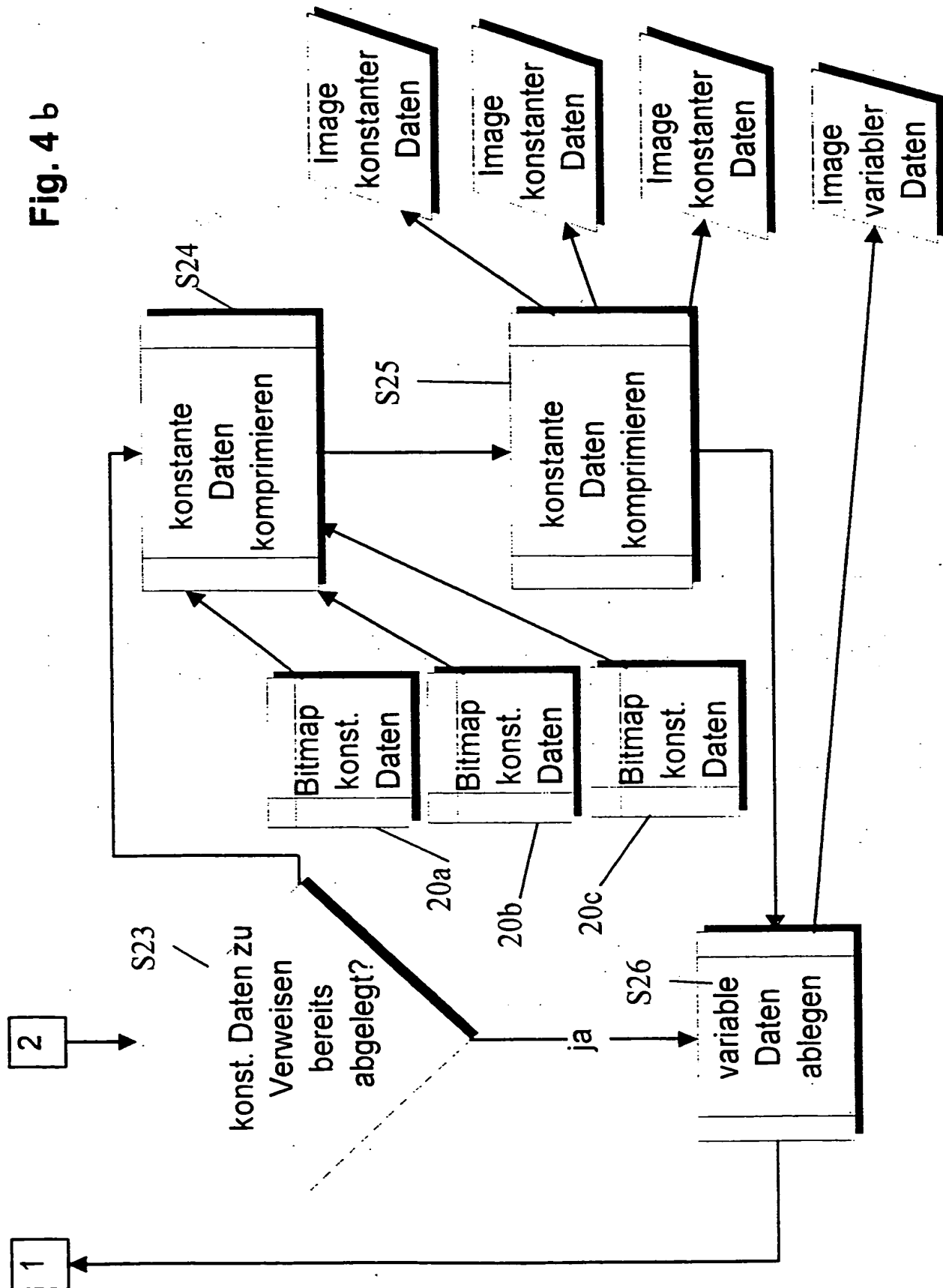
Fig. 4 a



**THIS PAGE BLANK (USPTO)**

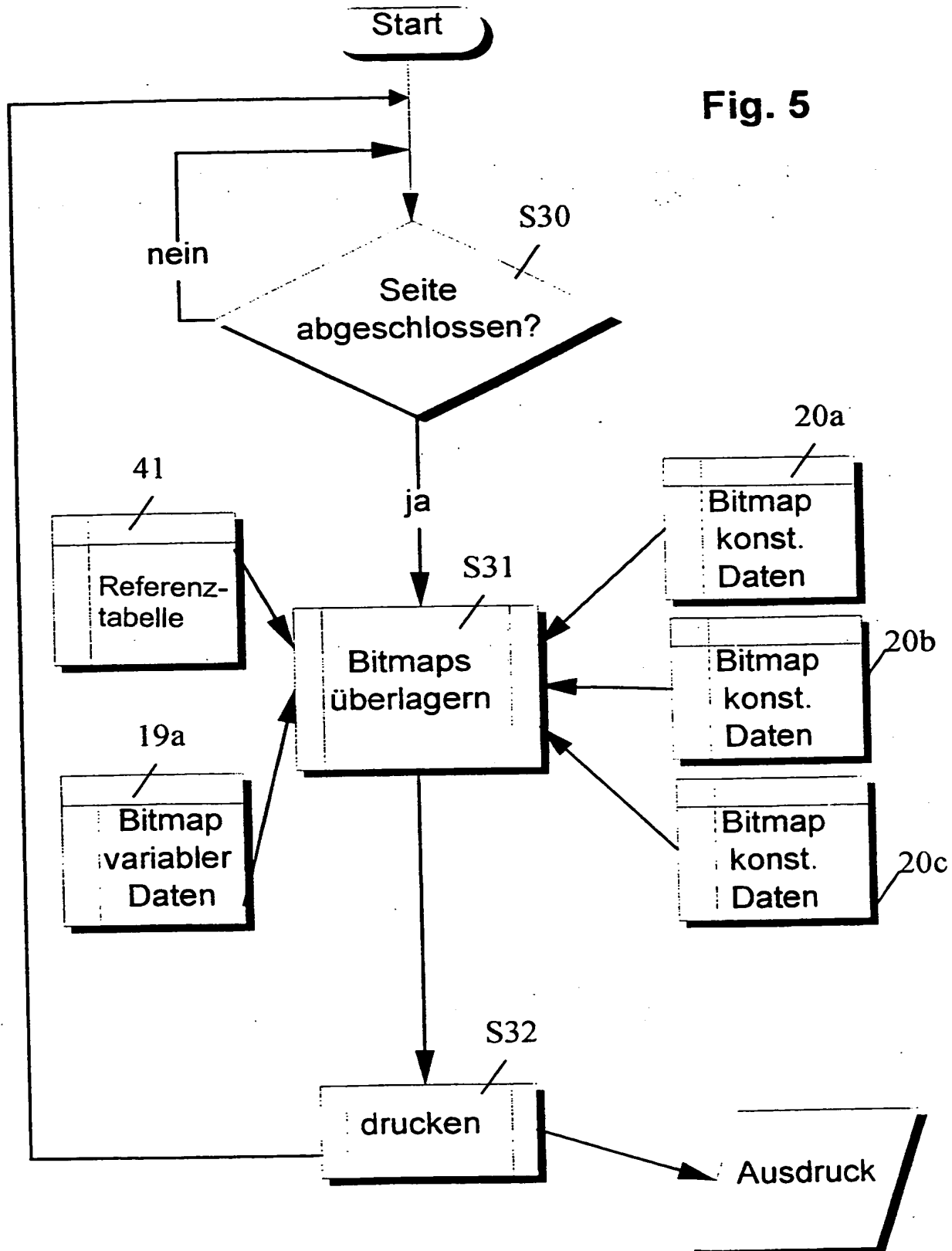


Fig. 4 b



**THIS PAGE BLANK (USPTO)**

Fig. 5



**THIS PAGE BLANK (USPTO)**

970803 10